Time Symmetry in Three Dimensions

Cristian López

University of Lausanne, Department of Philosophy
University of Buenos Aires, Institute of Philosophy
CH-1015 Lausanne, Switzerland
lopez.cristian1987@gmail.com

(*Draft June 11, 2020*)

Abstract

The notion of time reversal has caused some recent controversy in philosophy of physics. In this paper, I claim that the notion is more complex than usually thought. In particular, I contend that any account of time reversal presupposes, explicitly or implicitly, an answer to the following questions: (a) What is time-reversal symmetry predicated of? (b) What sorts of transformations should time reversal perform, and upon what? (c) What role does time-reversal symmetry play in physical theories? Each dimension, I argue, not only admits divergent answers, but also opens a dimension of analysis that feeds the complexity of time reversal: modal, metaphysical, and heuristic, respectively. The comprehension of this multi-dimensionality, I conclude, shows how philosophically rich the notion of time reversal is in philosophy of physics

Keywords: Time Reversal, Symmetry, Physical Laws, Modality, Time

1. Introduction

Symmetries play a paramount role not only in physical theories, but also in facing many fundamental philosophical problems in philosophy of physics and metaphysics (Nozick 2001, Baker 2010, Dasgupta 2016). Time-reversal symmetry, in particular, has increasingly drawn physicists' and philosophers' attention for mainly two reasons. First, it is relevant to address some long-standing philosophical queries around the nature of time (for instance, whether it manifests a privileged direction or not, see Horwich 1987, Price 1996, Arntzenius 1997). Second, it has proved to be crucial in empirical research as well as in guiding theory construction (Sachs 1987, Kleinknecht 2003, Sozzi 2008, Caulton 2016), which highlights its *explanatory* value in the foundations of physics. In addition, a recent philosophical discussion focuses on the meaning of time reversal in different physical theories, how it must be interpreted and defined in general and in particular cases (see, for instance, Savitt 1996, Albert 2000, Callender 2000, Earman 2004, Malament 2004, North 2008, Roberts 2017, Lopez 2019 among many others).

Regardless of how fruitful such literature has been for philosophy and foundations of physics, it generally promotes a one-dimensional angle in its analysis, as if time reversal were a simple notion. The literature puts the focus *either* on formal features (for instance, whether time reversal is implemented by an anti-unitary or unitary operator in quantum theories), on metaphysical aspects (for instance, whether it must change basic properties or not), or on heuristic virtues of time-reversal invariance (for instance, whether time-reversal invariance is an a priori symmetry or not), but it has failed to relate and articulate all these aspects. The aim of this paper is to defend that time reversal is a complex, multi-dimensional notion, where each dimension leaves room for opposing views with respect to time reversal in physics.

The multi-dimensionality of time reversal, I suggest, comes out of three questions that any approach to time reversal must answer:

(i) What is time-reversal symmetry predicated of?

(ii) What sorts of transformations should time reversal perform, and upon what?

(iii) What role does time-reversal symmetry play in physical theories?

Each question opens up a dimension of analysis that feeds the complexity of time reversal. I will contend that the first question involves a *modal dimension* since it leads to a distinction between a symmetry of the laws versus a symmetry of a solution (Section 2). The second question shows a *metaphysical dimension* since it leads to paying attention to the metaphysics of time underlying the notion of time reversal (Section 3). Finally, the last question proposes a *heuristic* and *epistemic dimension* since it deals with two opposing views on symmetries in physics –a by stipulation and a by-discovery view (Section 4). Each question will admit alternative answers, thus steering our understanding of time reversal towards different directions.

2. Modal Dimension: Symmetry of Solutions *versus* Symmetry of the Laws

The first question of this multi-dimensional analysis of time reversal is:

What is time-reversal symmetry predicated of?

In physics, it is common to speak of the symmetry group of a theory. For instance, it is said that non-relativistic quantum mechanics is Galilei invariant since it is invariant under the transformations of the Galilei group. Relativistic quantum mechanics is non-invariant under the Galilei group, but it turns out to be invariant under the Poincaré group. Similarly, it is common to speak of the symmetries of a theory when the symmetry transformation does not belong to a theory's group. Time reversal is a paradigmatic case of it, but there are many others (i.e. parity). In both cases, a symmetry is predicated *of* a theory, but in particular, *of* its dynamical equations –

If a physical theory remains symmetric under some transformation, this is so because its laws come out invariant under such a transformation. In this sense, a symmetry is a property of the laws.

However, it is also common to enquiry whether a specific model (or a specific system) is symmetric under a given transformation. In this sense, it is said, for instance, that a certain Lagrangian is symmetric under a symmetry transformation if it does not change the value of the Lagrangian. A single degree of freedom with Lagrangian

$$L = \frac{1}{2}\dot{q}^2 \tag{1}$$

is invariant under a transformation that shifts the coordinate q by an amount of δ, $S: q \rightarrow q + \delta$. So, in this case, we can assert that *this* Lagrangian is spatial-translation invariant. However, if we consider a slightly more complex system, for instance the trajectory of a particle through an inhomogeneous field, then the value of the Lagrangian will surely change as q is shifted. Therefore, this second Lagrangian will be non-spatial-translation invariant. In this way of speaking, a symmetry is predicated *of a model* (or *of a solution*) of a theory's dynamical equations; models that not only involve dynamical features, but also some non-nomic ones (for instance, boundary conditions, approximations and so forth). One point to stress here is that a failure of invariance for a specific model could be caused either by a theory's dynamics or by some non-nomic feature (for instance, a very special initial condition).

This ambiguity when it comes to the subject of a symmetry predication is captured by a distinction neatly drawn by Katherine Brading and Elena Castellani between *symmetries of solutions* (or of models) and *symmetries of laws*. In their words:

> "we must distinguish between symmetries of objects versus symmetries of laws (…). It is one thing to ask about the geometric symmetries of certain objects (…) and the asymmetries of objects (…). It is another thing to ask about the symmetries of the laws governing the time-evolution of those objects (…). Re-phrasing the same point, we should distinguish between symmetries of states or solutions, versus symmetries of laws" (2003: 1381).

The distinction remains valid for time reversal: we can say either that a dynamical equation is time-reversal invariant, or that a specific model of such an equation is. In my view, this difference is about their *modal* scope: when time-reversal invariance is predicated of a dynamical equation the assertion is meant to be *modally* stronger for it involves what is or isn't physically possible within a theory. When it is rather predicated of a specific model, its modal scope is much weaker because the (a)symmetry generally rest upon particular features of *that* model. In introducing a modal dimension in the notion of time reversal, the distinction is relevant also for some philosophical debates that crucially depend on the notion of time reversal –i.e. the debate around the arrow of time.

To see the distinction clearly, consider an analogy with Huw Price's intuitive corkscrew model (1996). Price proposes us to imagine a factory which generates equal number of left-handed and right-handed corkscrews. There are two points upon which we may want to focus. The first one is whether the production of corkscrews is symmetric. The second, whether an individual corkscrew is (spatially) symmetric or not. Price's model aims to show that, as there is no mystery whatsoever in having a symmetric production of left-handed and right-handed while having spatially asymmetric individual corkscrews, there is not mystery either in having time-reversal invariant laws and asymmetric models (see Price 1996: 88, 96). Otherwise stated, the mystery fades away when it is left clear we are dealing with two distinct sorts of questions. In particular, the difference resides in that we are attributing the predicate 'being time-reversal symmetric' to two different items. In the first case, we predicate 'time-reversal invariance' of physical laws (mathematically represented, in general, by differential dynamical equations), whereas in the second case the predicate is attributed to an evolution of the laws (mathematically represented by a solution of a differential dynamical equation). So, the differences in symmetry predications goes hand-in-hand with differences between laws and their solutions. Let us get into more details.

Consider Hamiltonian classical mechanics. Any solution of Hamilton's equation can be represented by a time-parametrized curve in the phase space (Γ)

$$\mathcal{E} = \{s_t \in \Gamma : t_i \leq t \leq t_f\} \tag{1}$$

The class of trajectories allowed by Hamilton's equations represents the class of its possible worlds W (or models), that is, those worlds wherein the laws are (approximately) true. The question whether or not Hamilton's equations are time-reversal invariant is the question of whether the equations generate (or produce) two *symmetric* subclasses of solutions, say, W^f and W^b (or time-symmetric twins, see Castagnino and Lombardi 2009). In addition, we require that if the laws are time-reversal invariant, then there exists a transformation T that maps solutions in W^f onto solutions in W^b. The existence of the bijective map $\mathcal{E} \rightarrow^T \mathcal{E}^T$ then secures that W^b keeps some structural or empirical equivalence[1] with respect to W^f. It follows from this that for any evolution $\mathcal{E} \in \mathrm{W}^f$ there will be a time-reversed evolution $\mathcal{E}^T \in \mathrm{W}^b$

$$\mathcal{E}^T = \{Ts_{Tt} \in \Gamma : -t_f \leq t_n \leq -t_i\} \tag{2}$$

And here the analogy is: as corkscrew factories produce a symmetric amount of left-handed and right-handed corkscrews, dynamical laws generate pairs of time-symmetric evolution, that is, those $\mathcal{E} \in \mathrm{W}^f$ and those $\mathcal{E}^T \in \mathrm{W}^b$. What is crucial to remark is that there is a strong modal element in time reversal as a symmetry of the laws. What kind of trajectories are allowed by a theory's dynamics marks off what is physically possible and what is not. To put it differently, an assertion

[1] Though I will here refer simply to empirical or structural equivalence, I acknowledge that a purely formal relation does not exhaust the meaning of a symmetry and it needs to be given with further constraints. However, which such constraints are has been matter of some controversy (see Belot 2013 and Dasgupta 2016).

of time-reversal invariance (or a failure thereof) says something about the whole class of solutions of a theory's dynamics; it says that if an evolution $\mathcal{E}_k$ is obtained, then it is physically possible (or impossible) to obtain a time-reversed evolution structurally or empirically equivalent, $\mathcal{E}_k^T$. So time reversal as a symmetry of the laws says something about the temporality of all possible models of the theory.

However, this says nothing about whether a specific trajectory is symmetric under time reversal or not. This happens because an evolution also involves non-nomic elements that characterize the evolution (and, even more, identify it). Even though a trajectory may be generated by a time-reversal invariant law, the trajectory itself can be asymmetric under the reversion of the direction of time. Terms can be a bit confusing here and that is why some authors prefer to talk about symmetry of the laws (as detailed above) in terms of *time-reversal invariance*, and about symmetry of the solutions in terms of *reversibility* (Price 1996 seems to be suggesting an akin distinction, Castagnino and Lombardi 2009 are clearer about it). Be that as it may, non-nomic factors (as boundary conditions) matter in determining whether a specific evolution is symmetric or not and can well explain why the model is asymmetric despite being generated by a time symmetric dynamic. But note that any symmetry assertion in terms of a symmetry of a solution is only valid for *that* solution and cannot be extended to other possibilities. In this sense, a symmetry assertion in terms of a symmetry of a solution is much narrower in modal scope than a symmetry assertion related to laws. In sum, whereas time reversal as a symmetry of the laws relates to the modal structure of a physical theory inasmuch as it circumscribes what is temporally possible and what is not for a theory's dynamics, time reversal as a symmetry of a solution just characterizes the actual temporal properties of a particular world (or model), which cannot be extended further than it.

This distinction, which may seem to be merely linguistic at first sight, is actually significant from a philosophical point of view. Take, for instance, the problem of the arrow of time, which has largely relied upon the notion of time reversal. The modal distinction between time reversal as a symmetry of the laws and as a symmetry of a solution introduces a distinction between two qualitatively different arrows of time. As above, there are now two questions as to time reversal and the arrow of time. One is whether a theory's dynamics treats the past-to-future direction (conventionally, t^+) and the future-to-past direction (t^-) differently. This is a question aimed at the dynamical equations of a theory and should be replied by their formal and structural features. The second is whether a specific model of a theory (for instance, the history of our universe, or this particular Hamiltonian in an inhomogeneous field) can be equally described with time running either backward or forward. This question is not exclusively aimed at a theory's dynamics, but also at other non-nomic factors that play an essential role in characterizing the model. Any arrow of time (or any rejection of it) we can extract from both questions are modally different, because they relate to items of a theory that vary in modal scope.

The discussion of the arrow of time has largely involved the obscure relationship between thermodynamics and statistical classical mechanics. On the one hand, thermodynamical systems evidently exhibit temporal biases as their entropies *always* increase. On the other, their mechanical reformulations are unable to capture such a feature, and the time asymmetry can only be explained, at best, statistically. The reason for this is that the mechanical equations of statistical classical mechanics are time-reversal invariant (in either of its formulations), so that for any entropy-increasing model, we can obtain a time-reversed entropy-decreasing model. There are many arguments of this sort (commonly known as "reversible objections") casting doubts on any intended reduction of thermodynamics to statistical classical mechanics. But these arguments did much more as they have also casted doubts on the explanation of the arrow of time in a classical setting. Evidently, if the underlying dynamics of macroscopic phenomena is given by classical equation of motions and these turn out to be time-reversal invariant, there is a fundamental sense in which the classical world lacks a direction of time. So, the problem is now how to explain the evident time asymmetry in terms of a directionless dynamics.

We can rephrase this scenario in terms of the corkscrew model: the equations of statistical classical mechanics are the factories that produce pairs of entropy-increasing models and entropy-decreasing models. This follows from the fact that they are time-reversal invariant. Here it comes the first question: does classical statistical mechanics treat the past-to-future direction (conventionally, t^{+}) and the future-to-past direction (t^{-}) differently? No, it doesn't. Nothing in the dynamics allows us to break the balance between entropy-increasing and entropy-decreasing models. In this sense, the theory is modally committed to rejecting any *fundamental* arrow of time –it is always dynamically possible to bring up a time-reversed model that satisfies the classical mechanical equations of motion. This symmetry declaration is valid for any model: nothing in its inner dynamics will be responsible for a time asymmetry.

To get this straight think of an inverse scenario, one in which a theory's dynamics turns out non-time-reversal invariant. What such dynamics tells us is that the set of solutions is *either* W^{f} *or* W^{b}, but cannot be both. In this case, every model will exhibit a time asymmetry that comes from the non-existence of a structurally or empirically equivalent model, regardless its initial or boundary conditions, or any non-nomic factor that will characterize it. So, we can pompously say that in such a case we have a quite good reason to declare that an arrow of time is necessary for that theory as the set of its possible models (or worlds) is either W^{f} or W^{b}.

Let us come back to the time-reversal invariant statistical classical mechanics. Whereas no time asymmetry can come from its dynamics, it is blatant that some classical models are time asymmetric (in fact, the majority of them). Here it comes the second question: does *this* model (with all its features, both nomic and non-nomic) equally allow descriptions with time running forward and backward? If the answer is negative, where do such temporal asymmetries come then from? There might be many answers to this, but one that has been quite popular (and controversial in equal degree) among philosophers and physicists is the so-called Past Hypothesis (see Albert 2000, Ch. 4). In a nutshell, the Past Hypothesis postulates very special initial conditions that, with some further assumptions, can explain temporal asymmetries, recover our temporally biased epistemic access to evidence and delivers the right thermodynamical predictions. All this despite having time-reversal invariant laws (for details, see Albert 2000, Callender 2000, Loewer 2012; for discussion, see Price 1996, Earman 2006 and Wallace 2011). The trick is that any probability distribution at any point of an evolution must be conditionalized over the Past Hypothesis. But, it should be stressed that the Past Hypothesis provides an explanation of why a concrete model of a physical theory (model that can be the history of our universe) is time asymmetric, but it says nothing about other possible models allowed by the theory. In this particular case, we obtain an explanation of why a model is time asymmetric, but this explanation strongly relies on non-dynamical elements that characterizes this model, but that cannot be extended further than that. The arrow of time so obtained is modally weaker since it is an arrow of time stemmed from an *a*symmetry of a solution, not of a law.

Summing up. I have shown that time-reversal symmetry can be predicated of two different mathematical items, which stand for two different physical notions: dynamical equations (physical laws) or solutions (evolutions), replicating Brading and Castellani's distinction between symmetry of laws and symmetry of solutions in the case of time reversal. Besides their conceptual specificity, I remarked that both predications are modally different. Whereas a symmetry of laws relates to what is possible and what is not for a physical theory, since it concerns the whole set of solutions of its dynamics, a symmetry of solution is much more modally circumscribed as it greatly depends upon some inner non-nomic features of a model.

3. Metaphysical Dimension: Relationalism *versus* Substantivalism

Time reversal is intended to act upon time and reverse it. But

What sorts of transformations should time reversal perform, and upon what?

This question relates to a wider one: what do we mean by 'time'? To provide an answer to these questions, I claim, we should dig into an (overlooked) metaphysical dimension in our understanding of time reversal. The motivation of this inquiry is simple: our metaphysical understanding of time guides our conceptualization and modelling of time reversal –If we are said to invert the direction *of time*, our course of actions will be different depending on what time is. And in this sense the metaphysics of time comes first: It determines not only what *time* reversal *is* but also *upon what* it is supposed to act. It has been suggested elsewhere (Lopez 2019) that our metaphysics of time plays an active role in modelling the time-reversal operator in quantum mechanics. Here I want to suggest that this is not exclusively the case for quantum mechanics but generalizes to any conceptualization of time reversal in physics.

There are at least two metaphysical views on time in philosophy –relationalist (or reductionist) and substantivalist (primitivist) views. The substantivalist holds that time is something real that exists independently of events and things placed within it. In addition, the substantivalist defends that time should be considered primitive in one's fundamental ontology, being consequently irreducible to anything else. Alternatively, the relationalist supports the idea that time is an abstract notion stemmed from events, things and their relations. Otherwise state, there is nothing like time that it is not already in dynamical features of things or events. The relationalist thus promotes a reductionist view of time –any temporal predicate can be ultimately boiled down to physical predicates related to the things' changing. Let's see all this in more detail.

Even though substantivalism and relationalism come in many flavors, there are some distinctive features that identify them and distinguish one another. To make explicit these distinctive features will be enough for the purposes of the paper. To begin, substantivalists believe that time (or space-time) exists independently of events and material things. In consequence, time instantiates properties over and above the properties that events, or material things instantiate. From this rather plain description, two main substantivalist tenets can be drawn

S1 **A dualist ontology**. There are two types of primitives in the substantivalism's ontology: material things or events *and* time (or space-time).

S2 **Independence and irreducibility**. Time's properties do not depend upon, or cannot be reduced to, events' or material things' properties.

As substantivalism comes in many flavors, so does relationalism. The details are not relevant here, but it is sufficient to say that relationalist reduces the structure of time to the structure of change, which may be qualified differently. The relationalist lesson is that time does not exist in any relevant sense of the word over and above the physical world, and that any property ascribed to time is ultimately reducible to a property attributed to events or the material things' changing. As before, two tenets:

R1 **A monist ontology**. There are only events or material things in the world plus their (space) temporal relations.

R2 **Dependence and reductionism**. Time's properties are reducible to events' or material things' properties. In this sense, the former metaphysically depends upon the latter. Consequently, any temporal parlance is nothing but a parlance about change. In a slogan: Time is *nothing but* change.

I opened this section claiming that the metaphysics of time comes first since it determines not only what time reversal is but also upon what it is supposed to act. This rather abstract assertion can be now fleshed out in terms of these metaphysical frameworks. The point is that substantivalists and relationalists will diverge over what time reversal is because they diverge over what time is and over the place it takes up in one's ontology.

For the substantivalist, time reversal is primarily a reversion of that primitive entity we name 'time' and of its intrinsic property related to its directionality (see Maudlin 2002). There is no metaphysical basis in the substantivalist's framework for time reversal to mean something different than a reversion of the direction of time *itself* (see North 2008: 212). Neither is there any place for a *reductio* of time reversal to any other transformation. Whatever we mean by time reversal is to be metaphysically exhausted by representing a reversion of time itself –any maneuver that intends to circumvent such a principle can, rightfully to my mind, be rejected by the substantivalist in terms of a relationalist maneuver.

The next natural step is thus to provide a schema for a substantivalist representation of time reversal. To begin, if time reversal chiefly intends to reverse the direction of time, any time reversal transformation will then represent a transformation of that parameter standing for 'time'. This can be straightforwardly implemented by the usual transformation, $T: t \rightarrow -t$. But the controversy is not about whether such a transformation should be included or not, but how it must be interpreted and what its relevance is when it comes to understand time reversal. In the majority of physical theories, time is an external parameter (t), which, under a substantivalist reading, exists and instantiates properties that are independent of the physical bodies. So, such a transformation *encodes* to a good extent what (substantivalist) time reversal is –just a reversion of the direction of time itself. A further step may be to stipulate that such a transformation generates a series of subsidiary dynamical transformations, to wit, of those magnitudes that intrinsically depends upon the variable t. To be clear: a dynamical magnitude will transform under time reversal if and only if it holds an intrinsic dependence upon time (i.e. if the magnitude is a first time derivative). It is worth noting that which magnitudes will change its sign under time reversal is a theory-relative issue.

The nature of time reversal for the relationalist looks quite different. To begin with, the transformation $T: t \rightarrow -t$ should not be taken too seriously, as t is just an abstract parameter that does not stand for any item in the relationalist's ontology. In her view, $T: t \rightarrow -t$ is nothing but a reparameterization of the time coordinate. So what? Here the reductionist attitude comes into play: a relationalist metaphysics of time implies that any temporal predicate, as for instance any reference to the 'directionality of time', should not be taken literally as if there were some primitive entity exemplifying the property of directionality. Rather, it should be taken metaphorically –the 'directionality of time' boils down to the directionality of the change of a series of temporal relations held by their relata. In consequence, time reversal is just a metaphor, so to speak, of a more fundamental transformation. What the relationalist has to find out is what it is the right sort of transformation that *realizes* time reversal within a physical theory.

How should the relationalist pursue this task then? First and foremost, it must be left clear that the guiding idea is that what is really substantive in the understanding of time reversal is not the transformation of time itself (for it is metaphysically nothing), but the transformation of *change*. This suggests that 'time reversal' should be considered a linguistic "shortcut" standing for dynamically relevant transformations related to the change (or motion) of a system. The directionality of time is, hence, *nothing but* the directionality of change. A reversion of the directionality of time is, therefore, *nothing but* a reversion of the directionality of change. To put it in a slogan: time reversal is *nothing but* change reversal. Actually, this is the overarching concept grounding the physical justification and guiding various implementations of time reversal in physical theories: the formal representation ultimately ought to capture the idea of reversing the change (or, more specific, motion), whatever it comes to mean within a physical theory (see, for instance, Wigner 1932: 325, Gibson and Pollard 1976: 177, Ballentine 1998: 377).

All this already delivers a general impression of what a (relationalist) time-reversal transformation should do and act upon. Besides the unphysical reparameterization of t, reversing the direction of time will be reversing those magnitudes that play a dynamical relevant role in an evolution. In particular, reversing those magnitudes is such a way that can formally represent a physical system evolving backward. To be emphatic: it is not the case that time reversal (somehow) produces a change in the sign of some magnitudes as if they were some subsidiary effects of reversing time, but that time reversal *is* such specific transformations.

Two points are worthy of mention. First, the metaphysical dimension of time reversal is partially independent of its physical and formal implementation. By this I mean that both metaphysical stances are committed to different understandings of time reversal. Just that. This does not per se entail that each of them finds a straightforward implementation within a physical theory or yields the same results (i.e. both keeping the same equations and models invariant under time reversal). It may not be so. If any of these metaphysical views leads to an untenable implementation of time reversal within a physical theory, this might lead (or not) to revise one's metaphysics, but it does not refute the thesis that metaphysical commitments with respect to time ground (and guide) our understanding of time reversal. Neither does it immediately imply that there was ultimately just one correct view, because it was the only physically viable.

Second, as I mentioned above, both views might entail implementations of time reversal that transform equations and models differently, and thereby, that render the same equations invariant and non-invariant. This is, of course, a quite interesting result, which deserves further philosophical inquiry. However, whether or not two distinct time-reversal transformations disagree on whether a given equation is left invariant, it is not crucial for the point I want to stress. What it is crucial is how they justify the properties of the time reversal transformation, regardless whether they deliver the same result or not. Let us see a concrete example to shed light on this. Suppose on how time reversal transforms momentum in Newtonian classical mechanics. Momentum is canonically defined as

$$p = m.v \tag{3}$$

The question 'how momentum transforms under time reversal', I claim, starts by firstly making explicit what we mean by time. So, we may put two alternatives on the table –either we are asking how momentum transform under relationalist time reversal or under substantivalist time reversal. If the former, we will seek the transformation that backtracks the Newtonian system to its original state, because time reversal *is nothing but* the action that generates such a backward evolution. So, we can rightfully declare that the trasformation

$$T : p \to -p \tag{4}$$

is part of the very definition of time reversal as it plays an essential role in generating the backward evolution. As it is part of the definition, it does not require further justification.

The substantivalist can declare that the (substantivalist) time-reversal transformation changes the momentum's sign, agreeing on the relationalist's result. But the *reasons* of why she would agree are different. For her, time reversal means a change of the direction of time itself, which is primarily given by

$$T : t \to -t \tag{5}$$

Now, she can *deduce* that, since momentum strongly depends on time as it definitionally depends on velocity (which is a first time derivative), time reversal *entails* the transformation of momentum's sign. This does not mean that time reversal be defined by the transformation of momentum's sign, but that a previously given definition of time reversal entails such a transformation. To be clear: the results are the same, but they are contingently the same in so much as the justifications of the results rest upon different bases. Naturally, that the results converge within a theory does not mean that they will converge in a different theory.

Summing up. Many of a time-reversal transformations' properties, I have argued, depend upon one's metaphysics with respect to time. This opens a metaphysical dimension in our understanding of time reversal in physics. In this section, I contrasted a relationalist-reductionist metaphysics of time with a substantivalist-primitive. The upshot was that both philosophical stances may disagree on what time reversal is and upon what it should act *because* they disagree on what time is. This dimension, I have claimed, chiefly concerns the justification to incorporate certain magnitude transformations either as definitions or effects of time reversal. In addition, this dimension is not idiosyncratic of concrete cases, but it pervades time reversal in physics.

4. Heuristic Dimension: Symmetries By-stipulation *versus* By-discovery.

In this section, I will address a third dimension of time reversal, which concerns some of its epistemic and heuristic aspects. Whereas the metaphysical dimension chiefly put the focus on the time-reversal *transformation* (what we metaphysically and physically mean by ‘time reversing’), this heuristic dimension rather centers in the status of *symmetries* in physics. To be precise, it centers in the epistemic and heuristic aspects that connect the construction of a time-reversal transformation to the role that the time-reversal *symmetry* should play in a physical theory. The question that triggers this analysis is

What role does time-reversal symmetry play in physical theories?

My claim is that this question can be answered from two opposing views with respect to symmetries in physics. One of them, which I will call *by-stipulation*, takes symmetries as postulated, being true independently of the details of the dynamics. The other, which I will call *by-discovery*, takes symmetries as a result of the details of the dynamics. In the former case, symmetries constrain the dynamics. In the latter, they depend on it. Both views, I contend, lead to a different understanding and modeling of time reversal in physical theories.

What does justify the distinction between by-stipulation and by-discovery symmetries? Brading and Castellani contend that some symmetry principles (mainly, space-time symmetries) seem to be used as *guides to theory construction*. That is, principles that must be satisfied whatever the final details of the theory come to be. The mechanism whereby a symmetry is raised to a principle that must be satisfied is that of *stipulation* –we postulate, independently of the details of a theory’s dynamics, that a given symmetry holds, and then that the dynamics must adapt to the symmetries’ constraints. When laying the groundwork for Bohmian Mechanics, Dettlef Dürr and Stephan Teufel for instance write

> “A symmetry can be a priori, i.e., the physical law is built in such a way that it respects that particular symmetry by construction. This is exemplified by spacetime symmetries, because spacetime is the theater in which the physical law acts (as long as spacetime is not subject to a law itself, as in general relativity, which we exclude from our considerations here), and must therefore respect the rules of the theater”. (2009: 43-44)

It is worth contrasting this quote to others we can find in the literature on symmetries. John Earman says

> "The received wisdom about the status of symmetry principles has it that one must confront a choice between the *a posteriori approach* (a.k.a. the bottom up approach) versus the *a priori approach* (a.k.a. the top down approach)". (2004: 1230)

Earman's distinction is in keeping with Brading and Castellani's (2007): whereas some take symmetries as *postulated*, guiding theory construction, others seem to follow an opposite trend, according to which symmetries are *a consequence of* specific laws –like a *discovery* (2007: 1347). The idea of *postulating* a symmetry entails certain degree of necessity for some aspects of a physical theory: its dynamics *must* satisfy the postulated symmetries come what may. Interestingly, Earman (1989) suggests that symmetry principles should be considered contingent, rather than necessary:

> "it would seem that the symmetry transformation could not fail to be a true symmetry of nature, contradicting the usual understanding that symmetry principles are contingent, that is, are true (or false) without being necessarily true (or false)" (1989: 121)

What these quotes make clear is that there are, implicitly or explicitly, at least two opposing approaches to the epistemic and heuristic status of symmetries in physics, involving time-reversal invariance as well. However, their characterizations look like a grab-bag of concepts. They resort on predicates like "necessary", "contingent", "a priori", "a posteriori", "being postulated or known before", "being discovered and known after", and so forth. It is thus not fully clear what the difference is truly about, so let us start by sorting things out.

To begin, the predicates "being necessary" and "being contingent" are *metaphysical*. These can be spelled out variedly, but one standard way to do it is by adopting the possible-world parlance,

a) x *is necessarily* Φ *if and only if* x *is* Φ *in every possible world wherein x exists;*

b) x *is contingently* Φ *if and only if* x *is* Φ *in some possible worlds, but it lacks* Φ *in others.*

The predicates "being a priori/posteriori" are rather *epistemic* –they are standardly defined in terms of whether something is known *independently of* experience. In which sense might a symmetry be a priori or a posteriori? If we take the terms strictly, that is, in relation to our experience, then the distinction does not make much sense for symmetries. Abstractly, a symmetry σ (as a symmetry of the laws, see Section 2) is a property of a mathematical structure $\mathfrak{E}$ –that is, a set of objects O equipped with relations R_i, and functions f_j, such that σ is an automorphism that maps O onto itself $\sigma: O \to O$ that preserves all of the relations and functions among objects in $\mathfrak{E}$. For physical theories, the implementation of this definition involves differential equations along with their space of solutions, and we say that a symmetry is a one-to-one mapping that preserves the space of its solutions (see Section 2). So, whether a dynamical equation is symmetric depends on the sort of formal relations held by the elements within it. And this is something we always know independently of the experience. Hence, it is always a priori in the strict sense. Symmetries may have an experimental manifestation, but this happens a fortiori and it is not independent of the theory (which, one way or another, already includes the symmetry in its dynamics). For instance, experimentally we might discover that a symmetry has been broken. But the symmetry itself is not derived from experience (see Healey 2009 for discussion). So, no workable epistemic distinction between a priori and a posteriori seems to apply.

I think, notwithstanding, that the a priori/a posteriori distinction could still be of some philosophical usage, if relaxed. The debate is not around whether symmetries are known before or independently of experience but known independently or before *the dynamics*. And so, I think Earman's and Dürr and Teufel's words should be understood. Also, this fits well with the idea that symmetries are either postulated or discovered, as Brading and Castellani put it. So, in this more liberal understanding of the distinction, we can say a symmetry σ is

(i) *A priori* if it is known independently of the dynamics of the theory.
(ii) *A posteriori* if its knowledge depends on the dynamics' details.

The predicates "being necessary" and "being contingent" now can be better interpreted as follows..

(a) σ is a *necessary* symmetry for T if there is no possible world wherein T is true without stipulating that T is σ-invariant.

If σ is postulated before the dynamics is given, then there is no possible world where such a dynamic does not satisfy that symmetry. In this sense, it is necessary: it is *required* by the very formulation of the dynamics –the symmetry is an essential property of the physical theory in so far as it was built under the assumption that the symmetry held.

(b) σ is *contingent* for T if there is at least one possible world wherein T turns out non-σ-invariant.

If symmetries are known (or discovered) depending on dynamics' details, nothing prima facie indicates that they must be considered as necessary for a physical theory –it can either have or lack symmetries. To be emphatic: there is no entailment between both kinds of predicates, but they are quite compatible when their meanings are relaxed. Both approaches can therefore be characterized as following:

By-stipulation approach	σ-symmetry must be regarded as *a priori* and *necessary* for a theory T
By-discovery approach	σ-symmetry must be regarded as *a posteriori* and *contingent* for a theory T

There are different justifications for taking either approach. For instance, Robert Sachs claims that a by-stipulation approach is required for the purpose of expressing explicitly the independence between the kinematics and the nature of the forces (Sachs 1987: 7). Others rely on a beforehand favored platonic view of time and space, which ideally are directionless (Dürr and Teufel 2009:

47). Alternatively, some just declare that time and space are not physically real, but belong to the geometrical framework from which we describe what is physically real (closer to any moderate o radical relationalist view). All these views directly or indirectly favor a by-stipulation approach – we are entitled to stipulate that time reversal must hold in some privileged cases, to wit, those where forces, fields and interactions vanish. From this stipulated time-reversal symmetric basis, we can understand "emerging" temporal asymmetries, largely due to forces', fields' and interactions' properties. Note that the modelization of time reversal must, therefore, tailor to such constraints since how dynamical magnitudes transform under time reversal are subjected to such constraints.

I think the reasoning the by-stipulation approach relies upon is the following. Suppose a free particle in Hamiltonian classical mechanics. If the equation that describes such a system turn out non-invariant under time reversal, the only responsible for the non-invariance would be the change of the chosen direction for the time coordinate. But the choice of a direction for the time coordinate is matter of convention, since it just says something about the perspective from which the system will be describe. But a change in our way to describe a system (or to write the equation down) should not induce a physical change. So, a free particle in Hamiltonian classical mechanics cannot be non-time-reversal invariant. We are therefore entitled to declare that free equations of motion must be time-reversal invariant. A by-discovery approach is not committed to such a consequence, but it will rather consider that there is no reason to stipulate that even free equation must be time-reversal invariant[2].

This heuristic and epistemic dimension is evident in the discussion on time reversal in classical electromagnetism, for instance. The following brief presentation will be enough to contrast both views in a simple case (for a good review of different positions, see Peterson 2015). On the one hand, David Albert (2000) has argued that classical electromagnetism is not time-reversal invariant because the Ampere's circuital law comes out non-time-reversal invariant. The explanation is the following. The magnetic field is a basic property of electromagnetism and basic properties (for ontological reasons) do not change sign under time reversal. It follows that, according to Albert, time reversal cannot change the sign of the magnetic field, so it is left invariant under time reversal. This fact renders Ampere's circuital law non-time-reversal invariant, since the right side keeps the positive sign, whereas a negative sign appears on the left. Albert's view is a good example of a by-discovery approach since time-reversal invariance is not stipulated, but it rather depends on the dynamics' detail and our ontological analysis of a theory's properties.

On the other, Frank Arntzenius and Hillary Greaves (2009) have brought up an alternative account that challenges Albert's –the "textbook account". According to it, the properties of time reversal do not depend on whether magnitudes are basic or non-basic, but on *postulating* that the

[2] Two points are worthy of mention. First, the by-stipulation approach seems to rely on a relationalist-like reasoning to justify the postulation of a symmetry. Second, the by-discovery approach seems to rather go along with an at least non-relationalist view since it can accept that in some situation free, fundamental equations of motion can be non-time-reversal invariant, leaving the door open for an asymmetry of time itself (whatever it comes to be conceptualized). This is nonetheless a quite speculative relation, which should be further argued.

(free, fundamental) equation is invariant and then figuring out the right transformation that keeps it invariant. In their words:

> "Next let us consider the electric and magnetic fields. How do they transform under time reversal? Well, the standard procedure is simply to assume that classical electromagnetism is invariant under time reversal. From this assumption of time reversal invariance of the theory (…) it is inferred that the electric field E is invariant under time reversal (…)" (Arntzenius and Greaves 2009: 6)

So, our ontological considerations regarding the magnetic field don't matter. Whether it changes sign under time reversal will depend on whether its sign's changing helps to keep the theory invariant or not. This is a clear case of the by-stipulation approach not only providing a view of how symmetries must be heuristically and epistemically regarded, but also guiding the modeling of the time-reversal transformation.

Summing up. Among philosophers and physicists there seem to be at least two opposing views regarding the status of symmetries in physics: by-stipulation and by-discovery. In this section, I have argued that the confrontation is also present in our characterization of time reversal, which not only affects conceptually how time reversal should be understood (either as a by-stipulation symmetry or by-discovery), but also guides the formal implementation of time reversal within a physical theory (for instance, in the case of electromagnetism).

5. Final Remarks

The aim of this paper was to bring to the fore the conceptual complexity associated with the notion of time reversal in physics, showing that it is not a simple, one-dimensional notion. I have argued that any conceptual elucidation of time reversal must give an answer to three questions:

(i) What is time-reversal symmetry predicated of?

(ii) What sorts of transformations should time reversal perform, and upon what?

(iii)What role does time-reversal symmetry play in physical theories?

These questions, I have shown, open three dimensions of analysis –modal, metaphysical, and heuristic, which admit of divergent views. This not only illustrates the philosophical richness of the concept of time reversal itself, but also shows how fertile is the notion of time reversal as terrain of philosophical inquiry about the nature of time and the status of symmetry in philosophy and physics.

6. References

Albert, D. Z. (2000). *Time and Chance*. Cambridge, MA: Harvard University Press.

Arntzenius, F. (1997). "Mirrors and the direction of time". *Philosophy of Science*, 64: 213-222.

Arntzenius, F. and Greaves, H. (2009). "Time reversal in classical electromagnetism". *The British Journal for the Philosophy of Science*, 60: 557-584

Baker, D. (2010). “Symmetry and the metaphysics of physics”. *Philosophy Compass* 5: 1157-1166

Ballentine, L. (1998). *Quantum Mechanics. A modern Development*. Singapur: World Scientific.

Belot, G. (2013). “Symmetry and equivalence”, in R. Batterman (ed.), *The Oxford Handbook of Philosophy of Physics*, Oxford: Oxford University Press.

Brading, K. and E. Castellani. (2007). “Symmetries and Invariances in Classical Physics”. *Handbook of the Philosophy of Science, Philosophy of Physics, Part B*. Eds. J. Butterfield and J. Earman. The Netherlands: Elsevier, 1331–1367

Callender, C. (2000). “Is time ‘handed’ in a quantum world?” *Proceedings of the Aristotelian Society*, 100: 247-269.

Castagnino, M. and Lombardi, O. (2009). “The global non-entropic arrow of time: from global geometrical asymmetry to local energy flow”. *Synthese*, 169: 1-25

Caulton, A. (2015). “The role of symmetry in interpretation of physical theories”. *Studies in History and Philosophy of Modern Physics*, 52: 153-162.

Dasgupta, S. (2016). “Symmetry as an epistemic notion (twice over)”. *British Journal for Philosophy of Science*, **67**: 837-878.

Dürr, D. and Teufel, S. (2009). *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory*, Berlin: Springer-Verlag.

Earman, J. (2004). “Laws, symmetry, and symmetry breaking: invariance, conservation principles, and objectivity.” *Philosophy of Science*, **71**: 1227-1241.

Earman, J. (2004a). “Laws, symmetry, and symmetry breaking: invariance, conservation principles, and objectivity.” *Philosophy of Science*, **71**: 1227-1241.

Earman, J. (2006). “The ‘past hypothesis’: not even false”, *Studies in History and Philosophy of Science*, 37: 399-430.

Gibson, W.M. and Pollard, B.R. (1976). *Symmetry Principles in elementary particle physics*. Cambrige: Cambridge University Press.

Horwich, P. (1987). *Asymmetries in Time*. Cambridge, MA: MIT Press.

Kleinknecht, K. (2003). *Uncovering CP violation*. Berlin: Springer-Verlag.

Loewer, B (2012). “Two accounts of laws and time”. *Philosophical Studies*, 160: 115-137.

Lopez, C. (2019). “Roads to the past: how to go and *not* to go backward in time in quantum theories”. *European Journal for Philosophy of Science*, 9: 27.

Malament, D. B. (2004). “On the time reversal invariance of classical electromagnetic theory”. *Studies in History and Philosophy of Modern Physics* 35, 295-315

Maudlin, T. (2002). “Remarks on the passing of time”, *Proceedings of the Aristotelian Society*, 102: 237-252.

North, J. (2008). “Two views on time reversal”. *Philosophy of Science*, 75: 201-223.

Nozick, R. (2001). *Invariances: The Structure of the Objective World*. Cambridge, MA: Harvard University Press.

Peterson, D. (2015). “Prospect for a new account of time reversal”. *Studies in History and Philosophy of Modern Physics*, 49: 42-56.

Price, H. (1996). *Time´s Arrow and Archimedes’ point: New Directions for the Physics of Time*. New York: Oxford University Press.

Roberts, B. (2017). “Three myths about time reversal invariance”. *Philosophy of Science,* 84, 2: 315-334.

Sachs, R. (1987). *The Physics of Time Reversal*. London: University Chicago Press.

Savitt, S. (1996). “The direction of time”. *The British Journal for the Philosophy of Science*, 47: 347-370.

Sozzi, M. (2008). *Discrete Symmetries and CP Violation: From Experiment to Theory*. Oxford: Oxford University Press.

Wigner, E. (1932). *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra.* New York: Academic Press (1959).